香港中文大學（深圳）
The Chinese University of Hong Kong, Shenzhen

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

IEEE SMC 2026
Bellevue, WA USA

A Unified Open-Set Framework for Scalable PUF-Based Authentication of Heterogeneous IoT Devices

Xin Wang[1], Peichun Hua[2], Chip Hong Chang[3], Wenye Liu[4] and Yue Zheng[1]

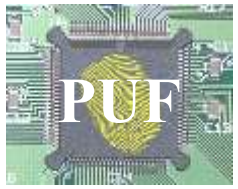
[1] School of Science and Engineering, The Chinese University of Hong Kong (Shenzhen)
[2] School of Data Science, The Chinese University of Hong Kong (Shenzhen)
[3] School of Electrical and Electronic Engineering, Nanyang Technological University
[4] Independent Researcher

- **Background: Why PUFs? Why Heterogeneity?**

Large-scale IoT fleets require lightweight, hardware-rooted authentication for resource-constrained devices.

Physical Unclonable Function (PUF)

Challenge
C

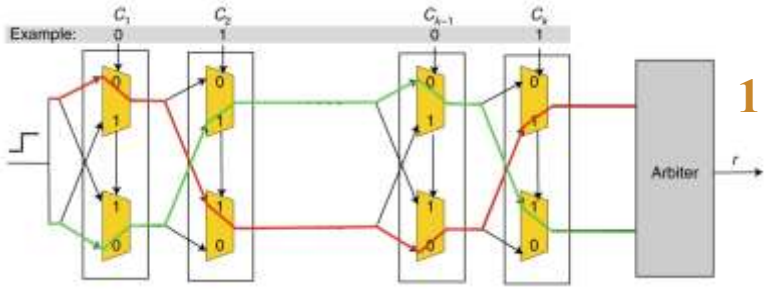


Response
R

On-the-fly generation

PUF exploits inherent manufacturing variations to generate device-specific responses to input challenges.

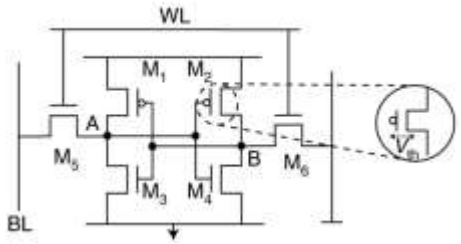
Strong PUF with exponential CRP space



1 bit per challenge

Arbiter PUF

Weak PUF with limited CRP space

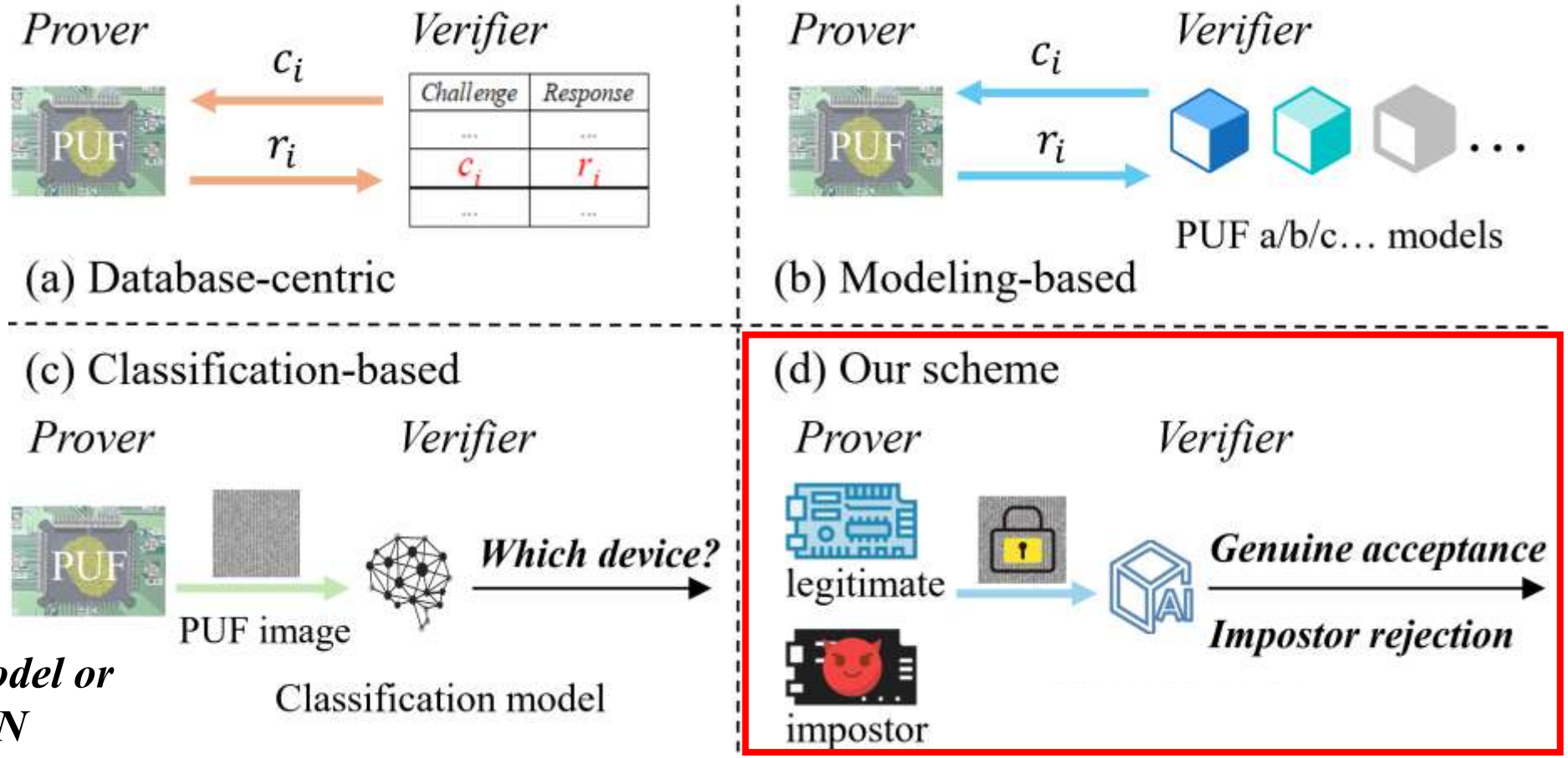




SRAM PUF

Multi-bits 2D array outputs

- **Prior Authentication Schemes and Our Approach**

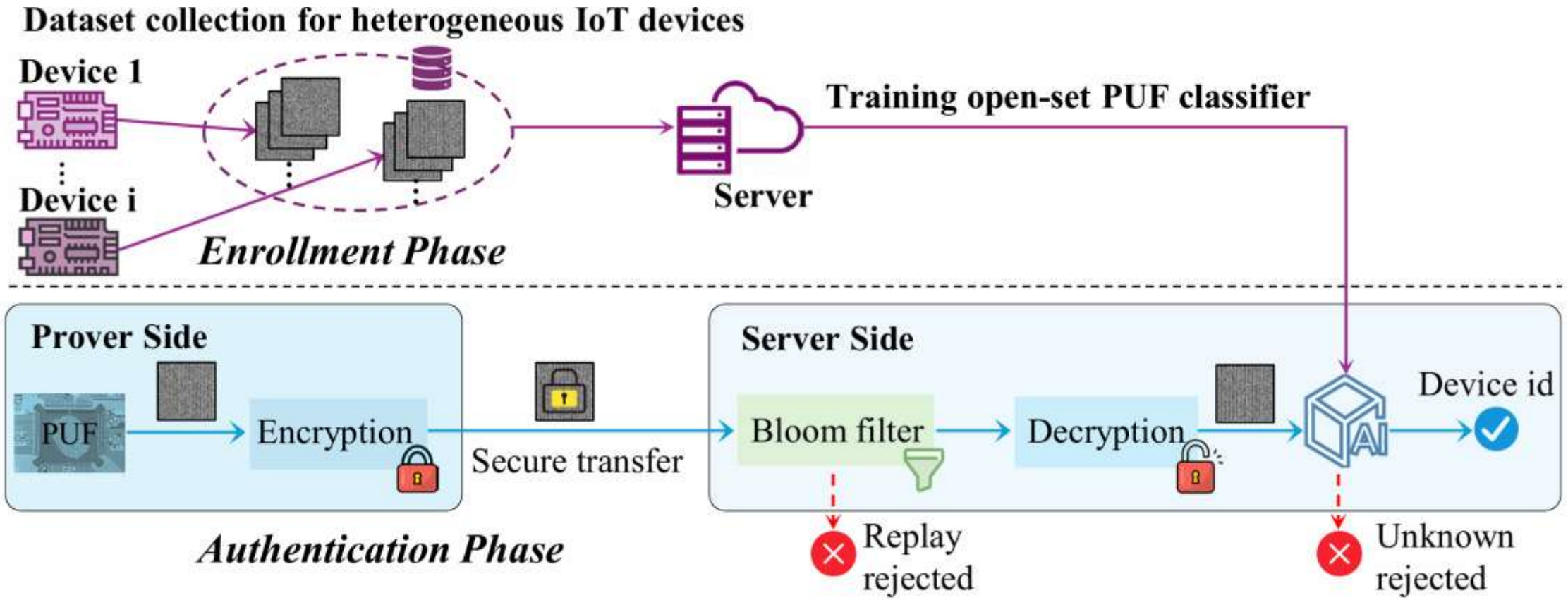


Closed-set model or Bayesian CNN

Open-set model

No demanding any error correction mechanism

- **Overview of Proposed Authentication Framework**



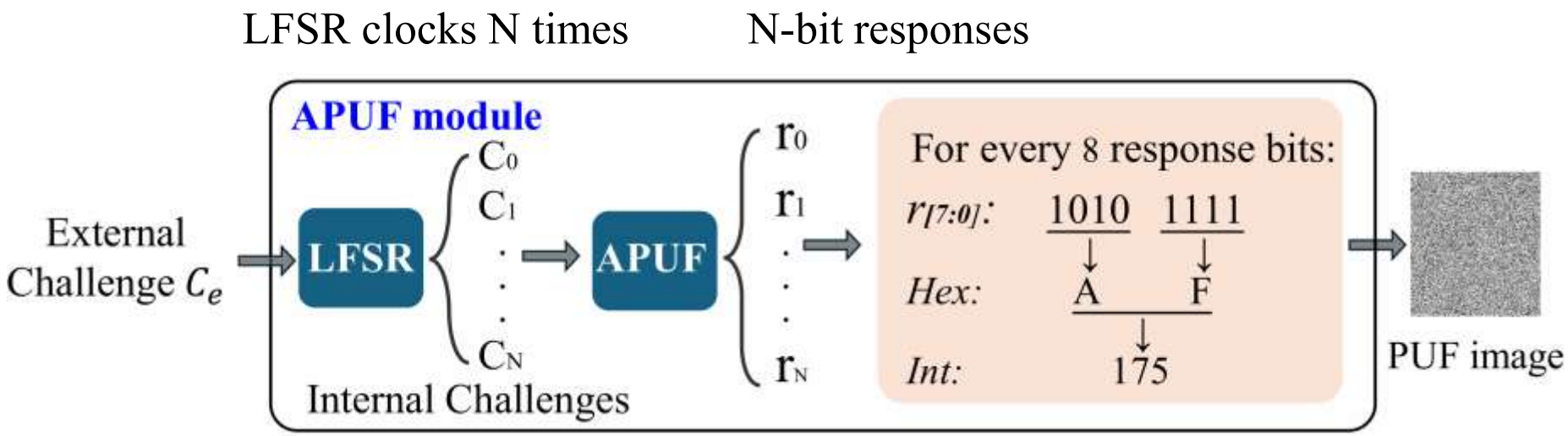
How can we obtain the device-specific PUF images for all PUF categories?

- **Unified Image Representation for Heterogeneous PUF Devices**

For memory PUFs: native 2D cell states (responses) → grayscale PUF images

For Arbiter PUFs: 1-bit response per challenge is incompatible with image-based classification.

Solution: LFSR (Linear feedback shift register) expands the initial challenge



- **Open-Set VS Closed-set for Authentication**

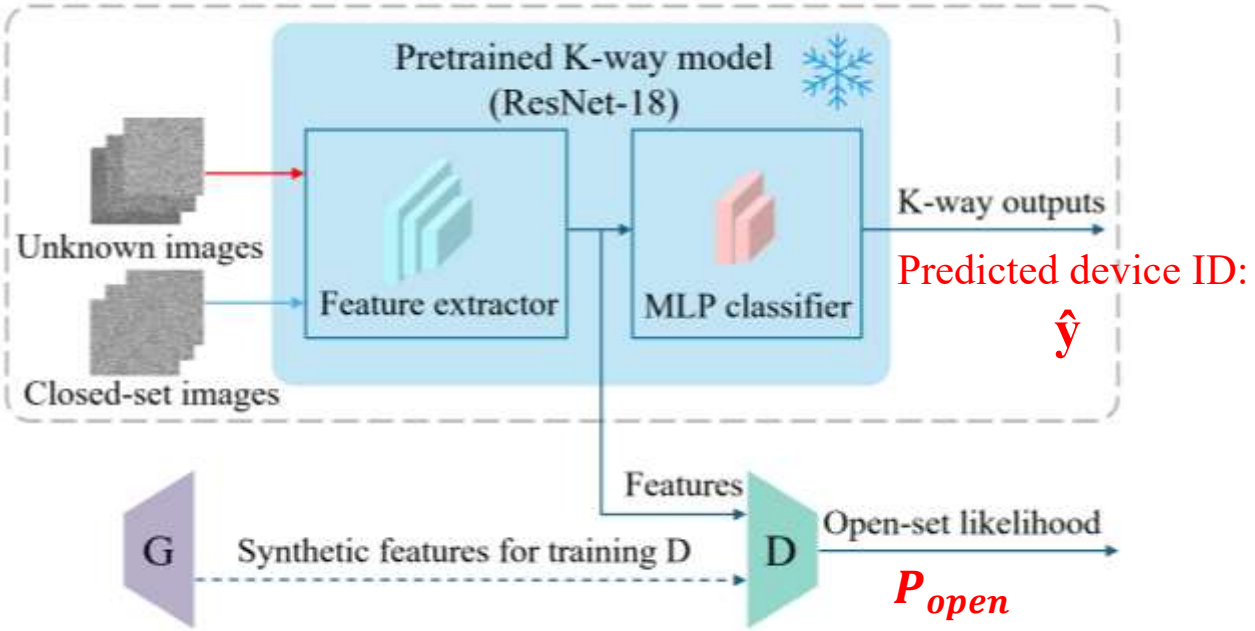
Closed-set: Always assigns an input to an enrolled device.
Open-set: Identifies enrolled devices while rejecting unknowns.

Two-stage Training

Stage 1 · Device identification:
Train ResNet-18 on enrolled-device images;
Freeze the feature extractor.

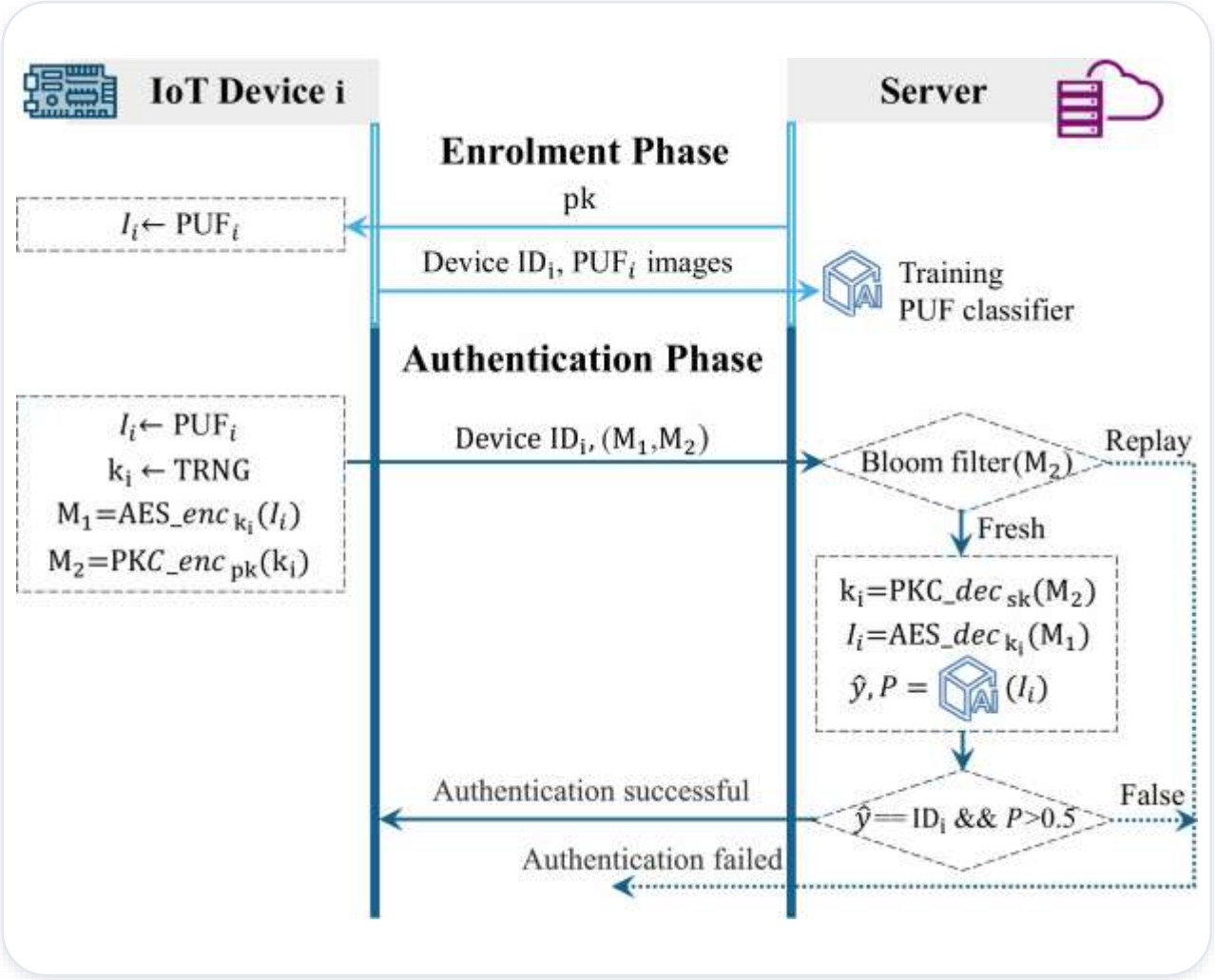
Stage 2 - Open-set training:
Generator G synthesizes features;
Discriminator D learns to distinguish enrolled from synthetic features.

Illustration of OpenGAN Architecture

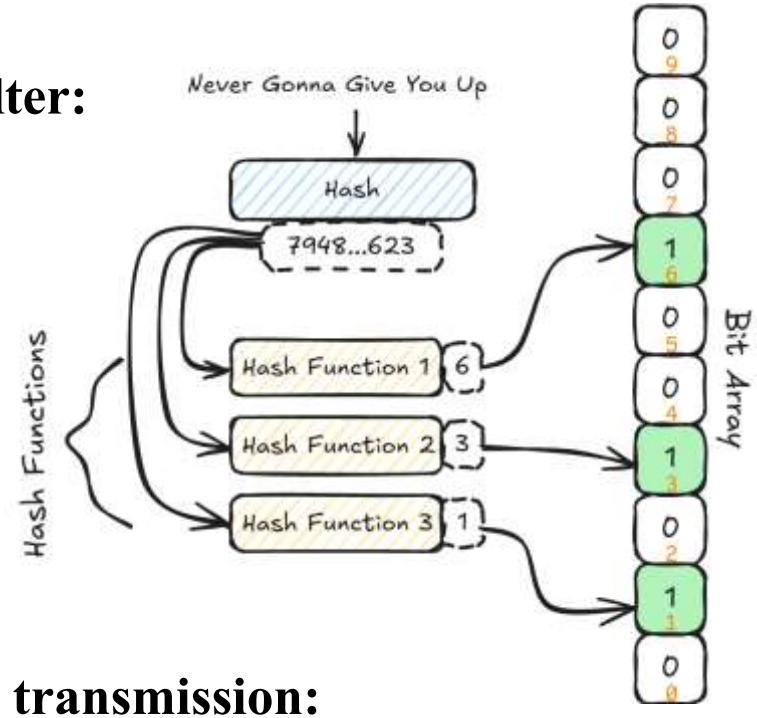


Accept iff $P_{open} > \tau$ and $\hat{y}$ = claimed ID.

- **Secure Authentication Protocol**



Bloom Filter:



Protected transmission:
AES-256-GCM protects the PUF image and detects tampering- RSA-OAEP encrypts the fresh session key.

Replay detection:
The server checks the encrypted session key (M_2) against a Bloom filter and previously seen M_2 values are rejected.

Encrypted transmission with replay checking before classification.

- **Experiment Settings: Datasets**

Realistic PUF response noise under environmental variations:

SRAM PUF — Response intra-device instability (average 5.85%, maximum 26.74%)

DRAM PUF datasets cover varying voltage (1.27–1.5 V) and temperature (20–50°C).

Four experimental settings for evaluating noise robustness and heterogeneous PUF authentication

Dataset	Train (legit)	Val (legit / imp.)	Test (legit / imp.)
APUF (ours)	25	25 / 6[a]	25 / 10[b]
DRAM PUF [26]	3	3 / 53[c]	3 / 53[c]
SRAM PUF [25]	40	40 / 52	40 / 52
Heterogeneous PUF	45	45 / 67[d]	45 / 68[e]

Impostors: (a) 4 APUFs + 2 DRAMs; (b) 7 APUFs + 3 DRAMs; (c) 1 DRAM + 52 SRAMs; (d) 8 APUFs + 59 SRAMs; (e) 8 APUFs + 60 SRAMs. Each legitimate device's images are split into train/val/test sets at a 3:1:1 ratio.

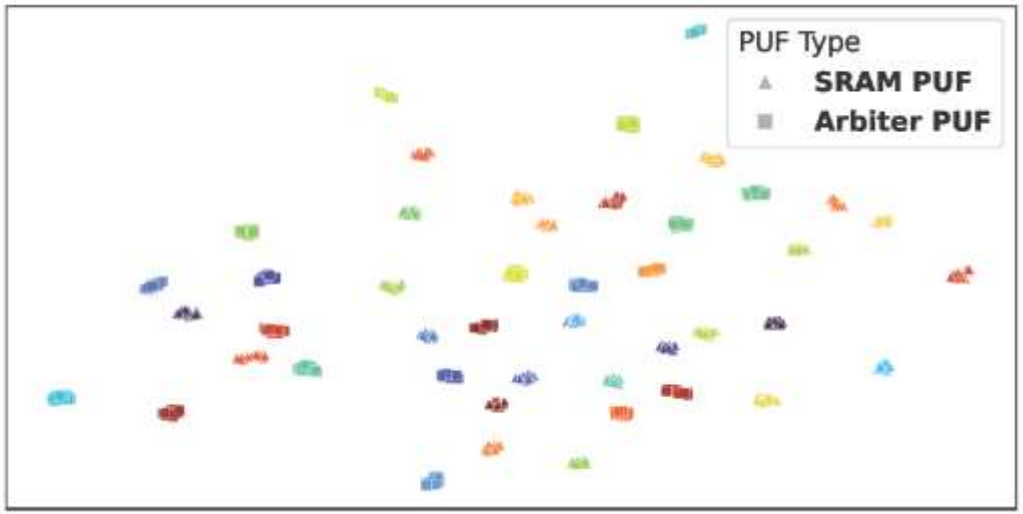
Hardware-disjoint impostors: Unknown devices are excluded from training and used only for validation and testing.

• Main Results

Method	Ours				Reproduced [28]	Mexis et al. [28]	Millwood et al. [26]	Karimian et al. [27]
PUF Type	Heterogeneous	Arbiter	SRAM	DRAM (noisy)	DRAM (noisy)	DRAM (nominal)	DRAM (noisy)	DRAM (nominal)
# Enrolled Devices	45	25	40	3	3	3	3-5	3
$Acc_{closed\text{-}set}$	100%	100%	100%	100%	100%	100%	92.7%-98.4%	98.6%
False Acceptance Rate	0.44%	0.35%	0.13%	0.33%	0%	0%*	✗	✗
False Rejection Rate	0.43%	0%	0.62%	2.04%	10.64%	0%*	✗	✗
# Test Samples (# Outlier Samples)	7784 (6859)	1440 (940)	6052 (5252)	5380 (5332)	5380 (5332)	172 (43)	116/154/192	108
Model Size	57.3 MB				40.4 MB	unknown	57MB	unknown
Authentication Time	≈ 0.57 s	≈ 0.67 s	≈ 0.57 s	≈ 0.64 s	Unknown	≈ 20 s	5.72–7.43 s	Unknown
Replay/MITM Attack Resistance	✓	✓	✓	✓	✗	✓	✗	✗

Key findings:

- 100% closed-set accuracy across all four evaluation settings
- 45 heterogeneous devices: 0.44% FAR and 0.43% FRR
- Distinct device clusters across both PUF types in the UMAP projection
- Authentication time: 0.57–0.67s, versus ~20s reported for BCNN



UMAP visualization for extracted device features. Color: device ID; shape: PUF type.

- **Ablation and Scalability Evaluation on SRAM PUFs**

TABLE III: Impact of image size on SRAM PUF dataset (40 devices). Mean ± std over 5 random seeds.

Size	FRR (%)	FAR (%)	AUROC	F1
25 × 25	2.40±1.85	1.84±1.15	0.9938±0.0102	0.9320±0.0409
50 × 50	**1.20±0.68**	**0.05±0.05**	**0.9993±0.0003**	**0.9927±0.0038**
100 × 100	6.90±3.75	0.51±0.86	0.9986±0.0005	0.9508±0.0210
200 × 200	18.47±3.76	4.18±3.16	0.8963±0.1814	0.5620±0.3892

50 × 50 gives the best overall results among the tested size.

TABLE IV: Impact of discriminator hidden dimension (n_d) on SRAM PUF dataset (image size: 50 × 50). Mean ± std over 5 random seeds.

n_d	FRR (%)	FAR (%)	AUROC	F1
64	8.00±2.20	6.24±3.99	0.9625±0.0186	0.8176±0.0843
128	**1.12±0.74**	0.46±0.43	0.9993±0.0008	0.9796±0.0109
256	1.2±0.68	0.05±0.05	**0.9993±0.0003**	**0.9927±0.0038**
384	18.33 ± 14.18	**0.00 ± 0.00**	0.9896 ± 0.0190	0.8936±0.1040

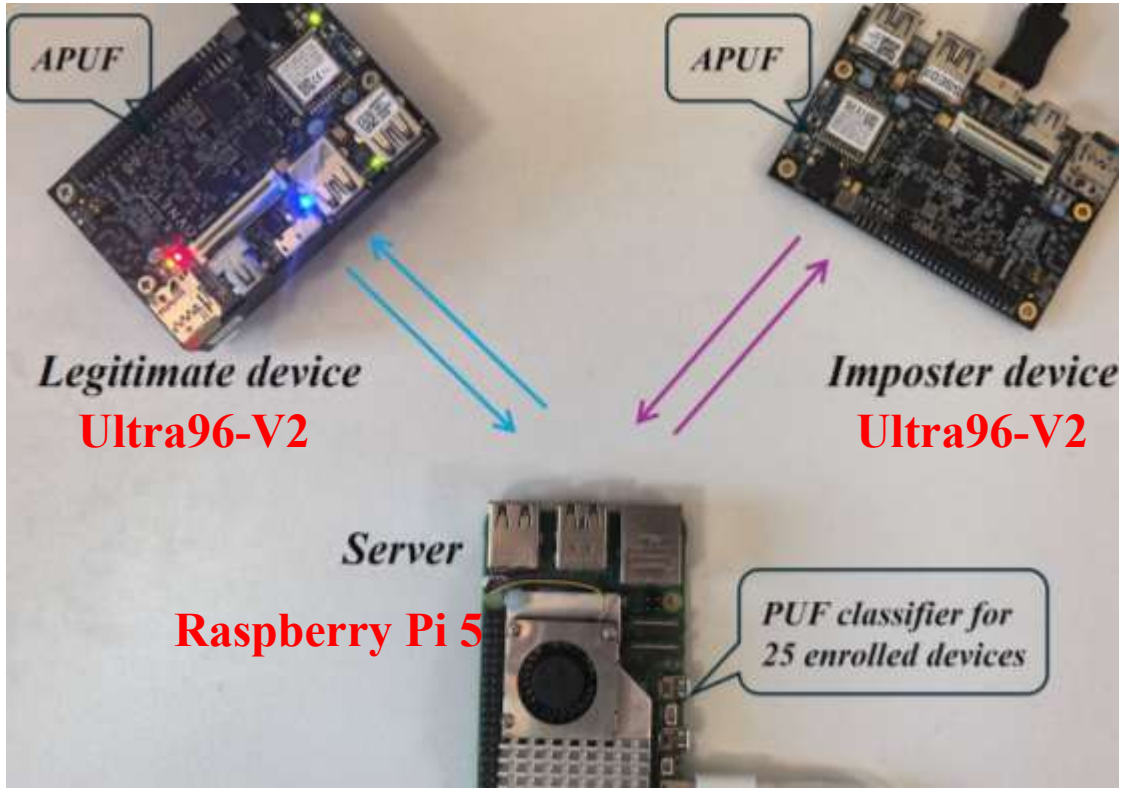
Hidden dimension 256 balances false acceptance and false rejection.

TABLE V: Scalability on SRAM PUF devices (image size: 50 × 50). Mean ± std over 5 random seeds.

#devices	FRR (%)	FAR (%)	AUROC	F1
5	0.20±0.45	0.38±0.22	1.0000±0.0000	0.9085±0.0477
10	0.90±0.65	0.31±0.60	0.9999±0.0001	0.9602±0.0657
20	0.70±0.33	0.01±0.01	0.9985±0.0007	0.9962±0.0015
30	0.73±0.49	0.03±0.06	0.9992±0.0004	0.9950±0.0037
40	1.20±0.68	0.05±0.05	0.9993±0.0003	0.9927±0.0038

Low FAR and FRR across 5–40 enrolled devices.

- **Hardware Prototype**



Configuration:

• Server: Raspberry Pi 5 (deploying classifier)

• Device: Two Ultra96-V2 FPGA boards implementing APUFs — one legitimate, one impostor

Key measurements:

• Authentication latency (encrypt + transmit + decrypt + infer) : ~0.67 s

• Communication overhead: 2,958 bytes per session (50×50 image, RSA+ AES)

• Model size: **57.3 MB**

• Classifier Inference cost: 0.25 GFLOPs per authentication

- **Conclusion**

  - A common image representation enables authentication across heterogeneous PUF types;
  - One open-set classifier identifies enrolled devices and rejects unknown devices without helper-data-based key reconstruction;
  - A lightweight secure protocol with replay detection is demonstrated on an embedded prototype.

Future work:

Larger fleets;

Additional PUF types;

Long-term working stability.

香港中文大學（深圳）
The Chinese University of Hong Kong, Shenzhen

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

IEEE SMC 2026
Bellevue, WA USA

Thank you for your attention!

PASSLab website

Contact

Xin Wang: xinwang3@link.cuhk.edu.cn
Prof. Yue Zheng: zhengyue@cuhk.edu.cn