barkhauseninstitut.org

A Multi-head CNN-based Side-channel Attack on the Energy-efficient DIZY Stream Cipher

George Attia

31.03.2026

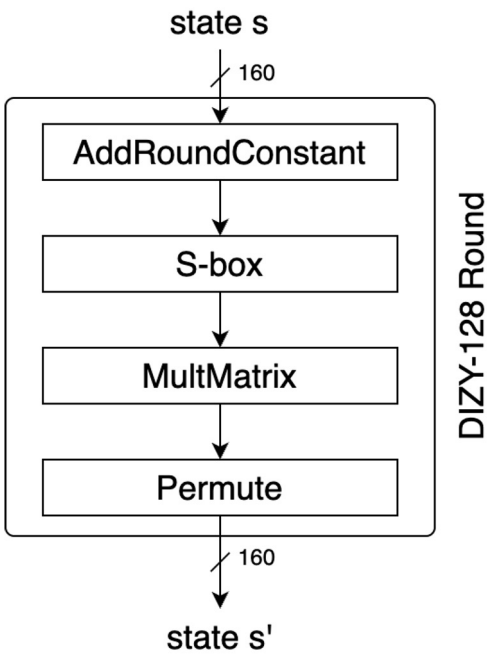
Introduction

- Side-Channel Attacks **(SCAs)** exploit physical leakages, such as power consumption during cipher execution.
- **Deep Learning (DL)-based** SCAs outperform traditional methods such as **CPA** by modeling complex leakage patterns.
- Previous research predominantly targets block ciphers such as AES. **Lightweight stream ciphers** remain understudied.
- We evaluate the feasibility of DL-based SCAs on hardware implementations of **DIZY** stream cipher as a case study.

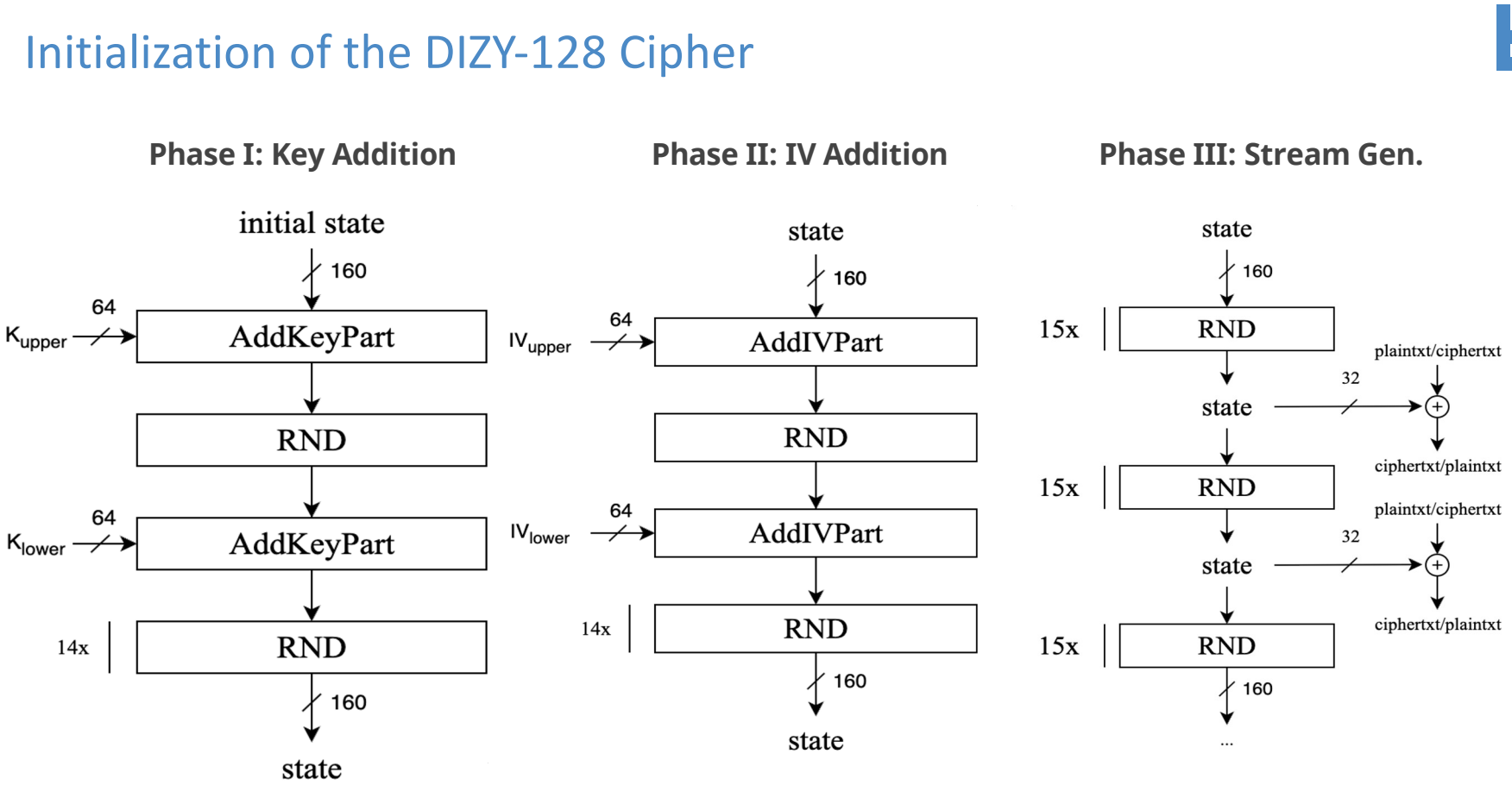
DIZY-128 Cipher: Round Operations

DIZY is a round-based cipher. Each round performs the following operations:

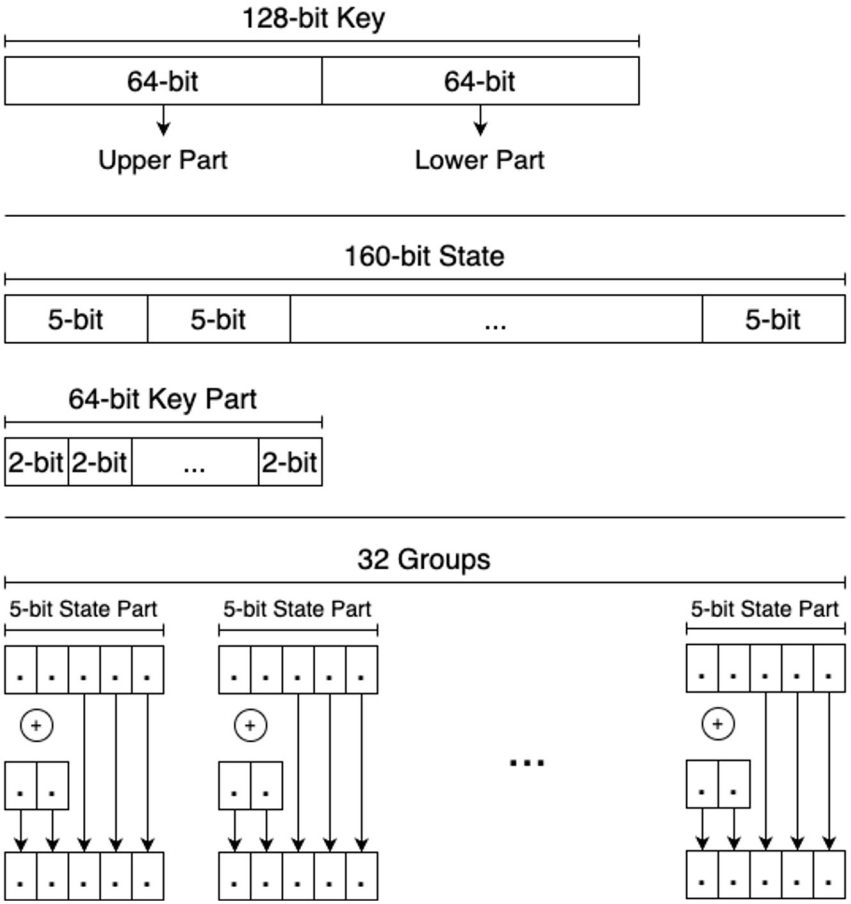
- **AddRoundConstant**
  - Defined using a 4-bit Linear-Feedback Shift Register (LFSR).
  - Adds a **4-bit constant** to each **5-bit group** in the state.
- **S-box**
  - **5-bit S-box.**
  - Applies the DIZY S-box to every **5-bit group** in the state.
- **MultMatrix**
- **Permute**



Initialization of the DIZY-128 Cipher

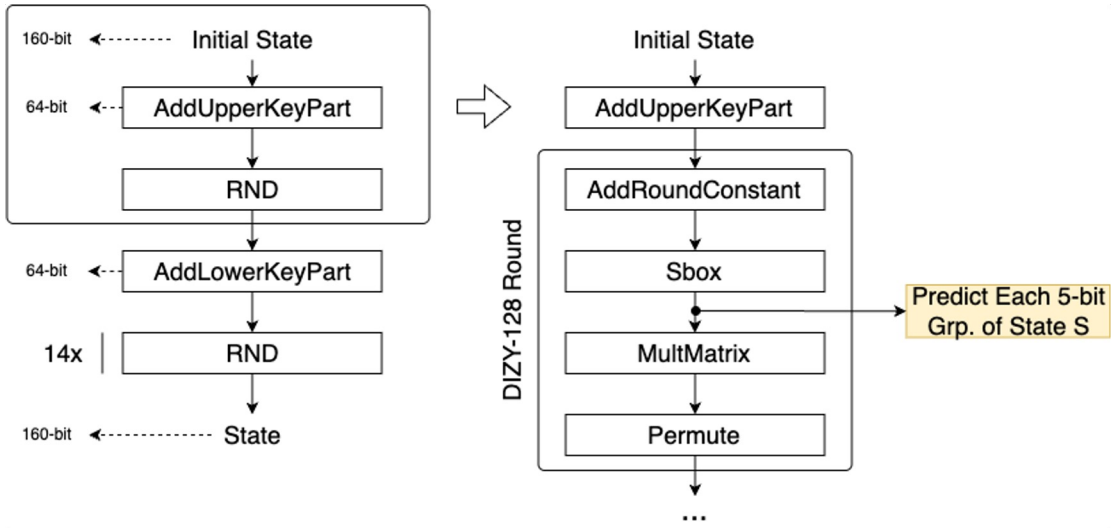


DIZY-128 Cipher: Add Key/IV Operation



DIZY-128 Cipher: Side-Channel Leakage Model

- Train a neural network to predict each 5-bit group of the output of the first S-box.
- Opting for 5-bit groups will result in a ML classification problem with **2^5 labels**.
- **The 64-bit upper key part** is expected to be revealed using this leakage model.



SCA Mathematical Formulation

- Let S_0 be the initial state, split into 32 groups of 5 bits.
- Let K be the 128-bit secret key. K_{upper} is the upper 64 bits split into 32 groups of 2 bits.
- Let C be the 4-bit constant used in the first round, generated by an LFSR.

1. Key Addition

$$G_{1,g} = G_{0,g} \oplus (K_g \lll 3)$$

2. Constant Addition

$$G_{2,g} = G_{1,g} \oplus C$$

3. S-box

$$G_{3,g} = \text{S-box}(G_{2,g})$$

- An ML model predicts

$$P_{g,v} = P(G_{3,g} = v)$$

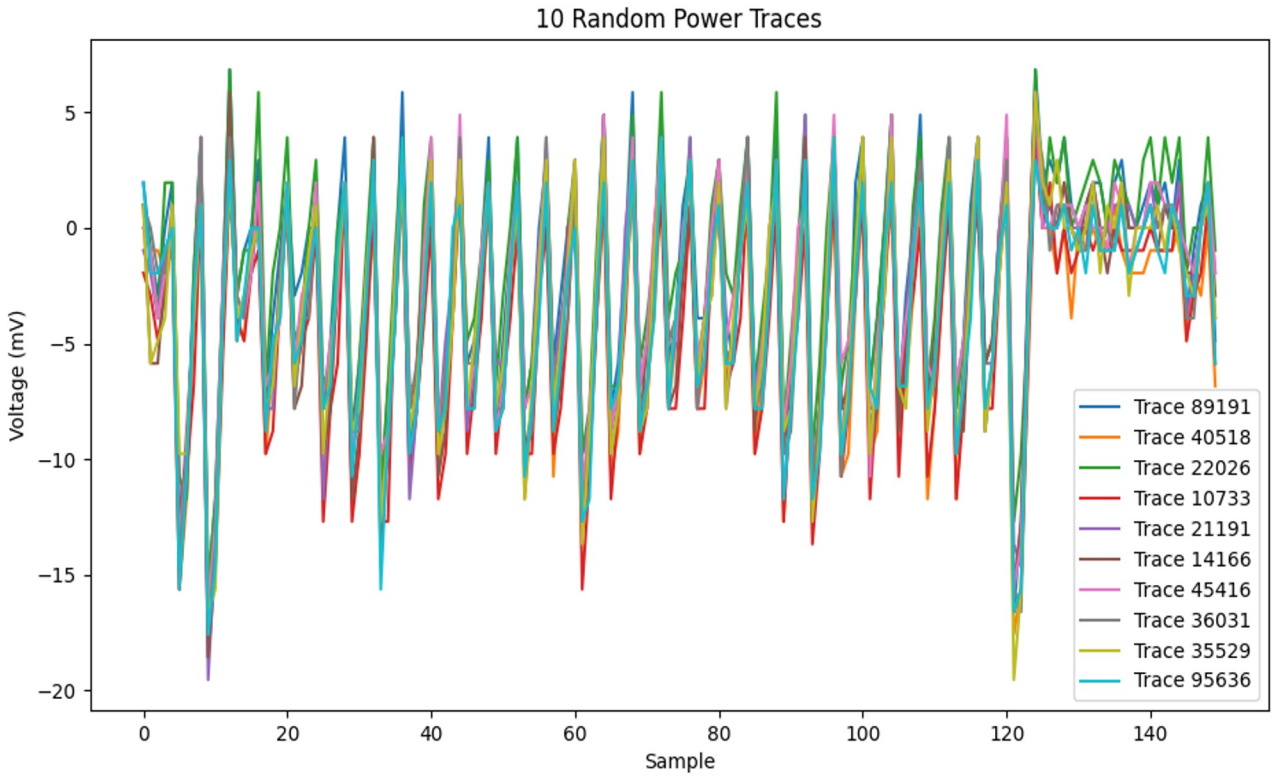
$$\text{for} \quad v = 0, 1, \ldots, 31$$

Collected Power Traces

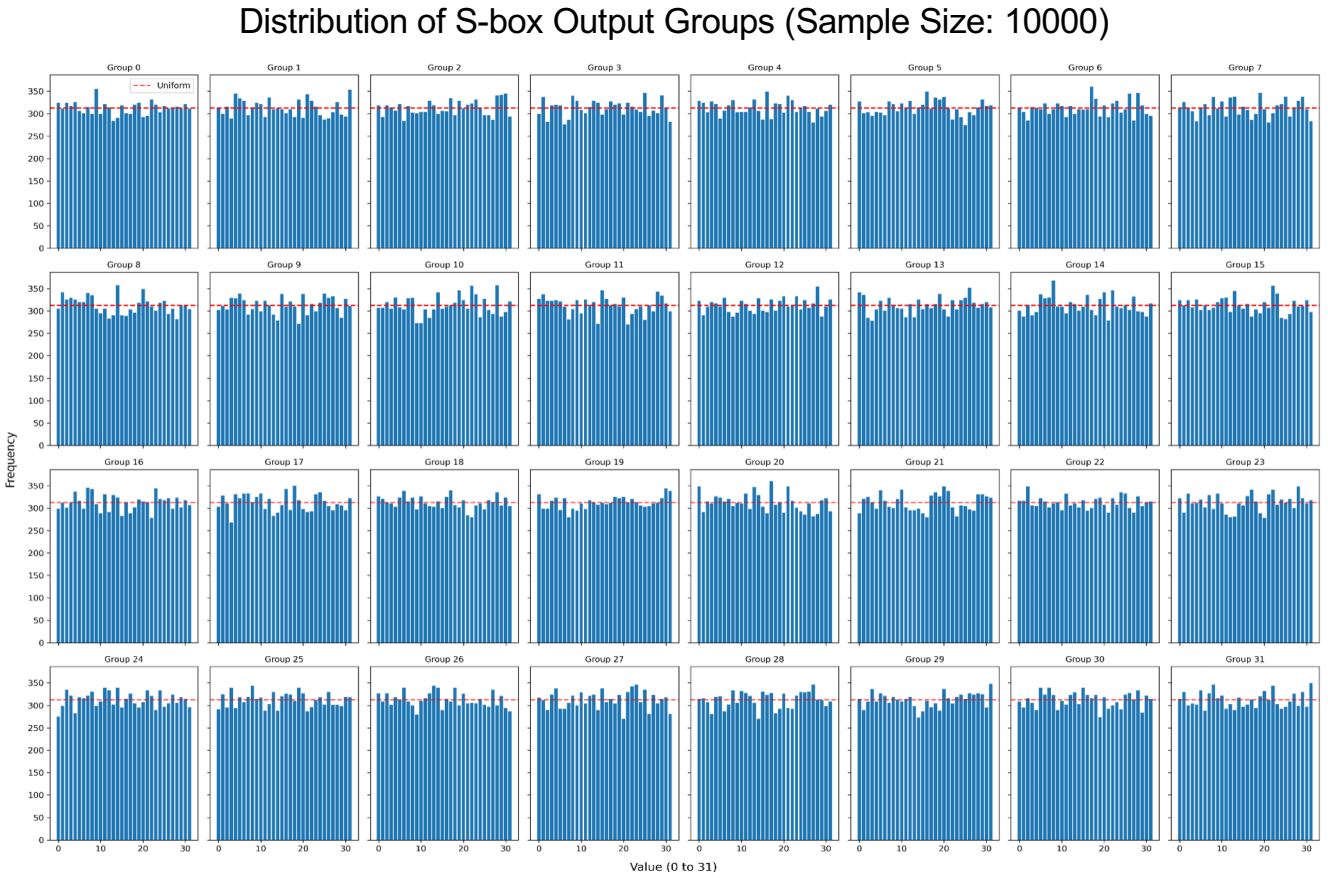
- Three power trace datasets were collected to train the deep learning models for the SCA.

Attribute	Fixed-Key Dataset	Variable-Key Dataset	Analysis Dataset
Total Number of Traces	250,000	500,000	500,000
Number of Keys	1	10	10
Traces per Key	250,000	50,000	50,000
Sampling Rate	4 samples/cc	4 samples/cc	4 samples/cc
Sampling Type	Synchronous	Synchronous	Synchronous

Collected Power Traces Sample

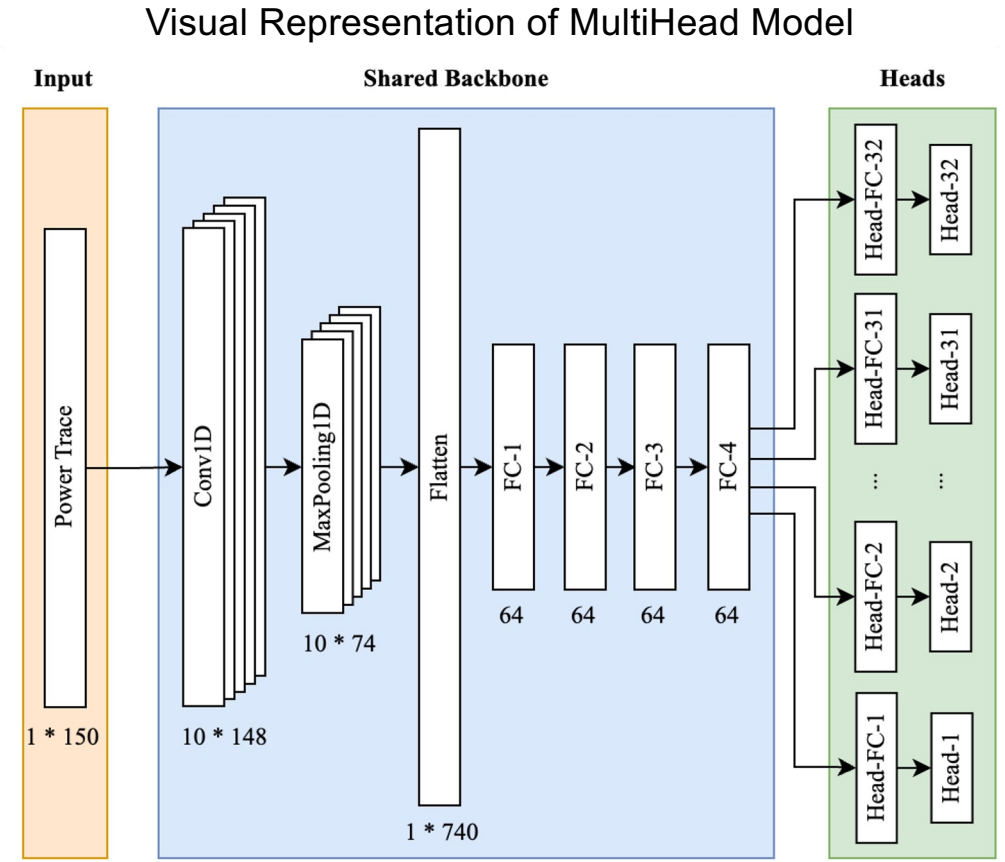


Label Distribution Analysis



CNN Model Architectures

- Two CNN architectures were developed to perform an SCA on the DIZY cipher
  - SeparateNetworks
  - MultiHead



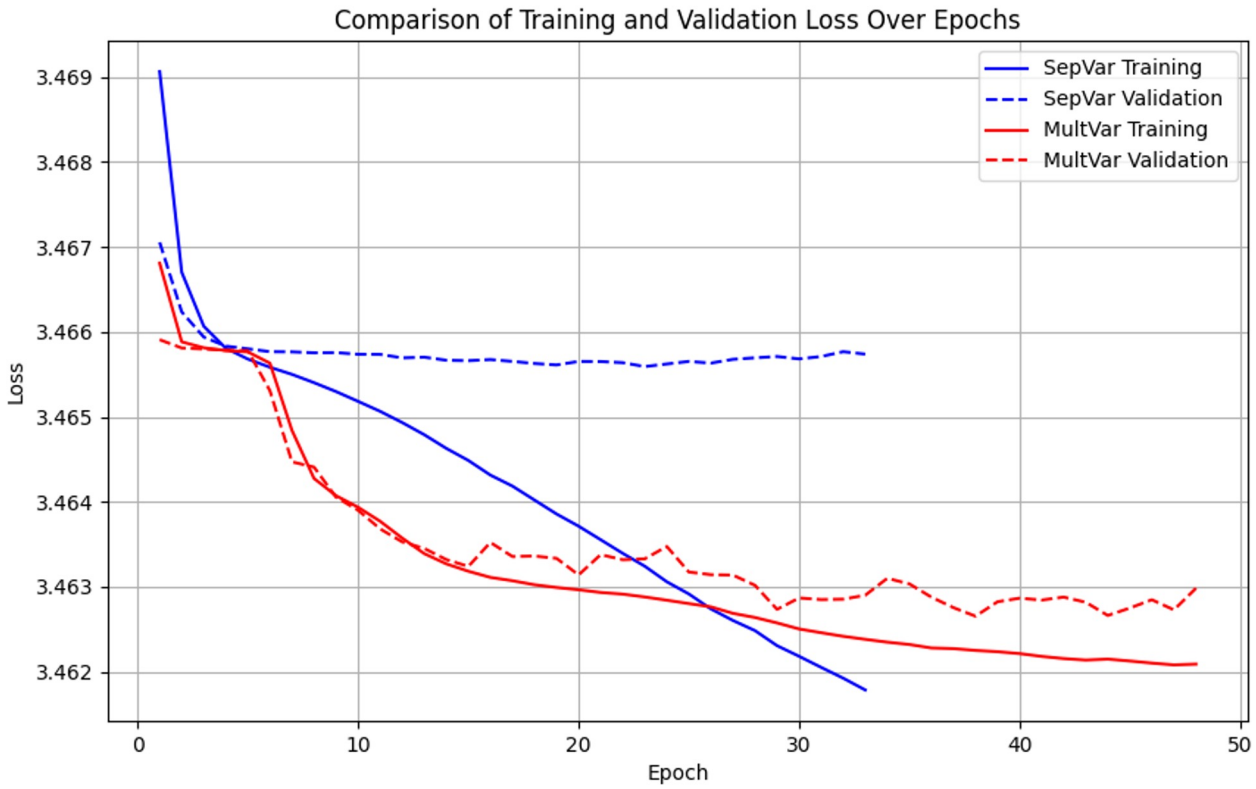
Training Experiments

- **Six CNN models** were trained to assess the success of the SCA.

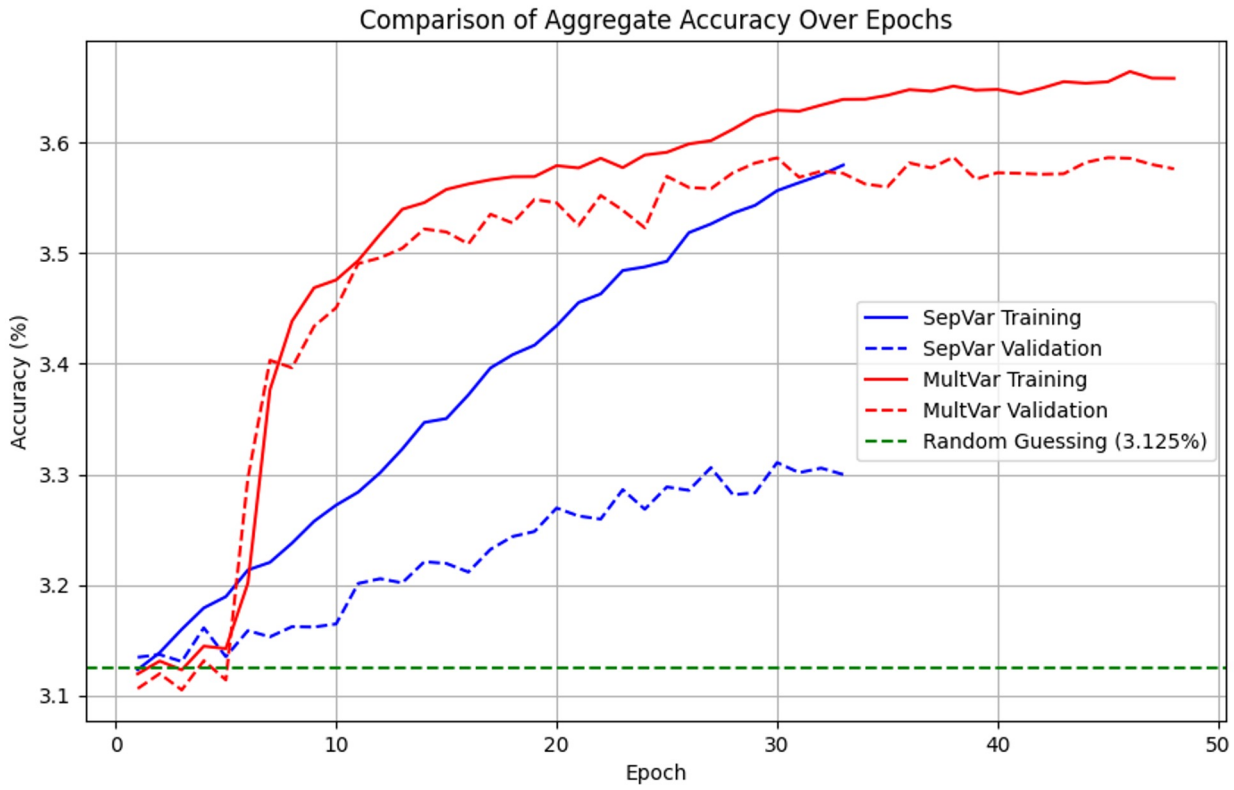
Summary of Trained Models

ID	Architecture	Dataset Configuration	Traces (Train/Val./Test)	Keys (Train/Val./Test)
SepS	SeparateNetworks	Fixed-Key, Small	90K/10K/50K	Same Key
MultS	MultiHead	Fixed-Key, Small	90K/10K/50K	Same Key
SepL	SeparateNetworks	Fixed-Key, Large	180K/20K/50K	Same Key
MultL	MultiHead	Fixed-Key, Large	180K/20K/50K	Same Key
SepVar	SeparateNetworks	Variable-Key	400K/50K/50K	8/1/1
MultVar	MultiHead	Variable-Key	400K/50K/50K	8/1/1

SepVar and MultVar - Loss over Epochs

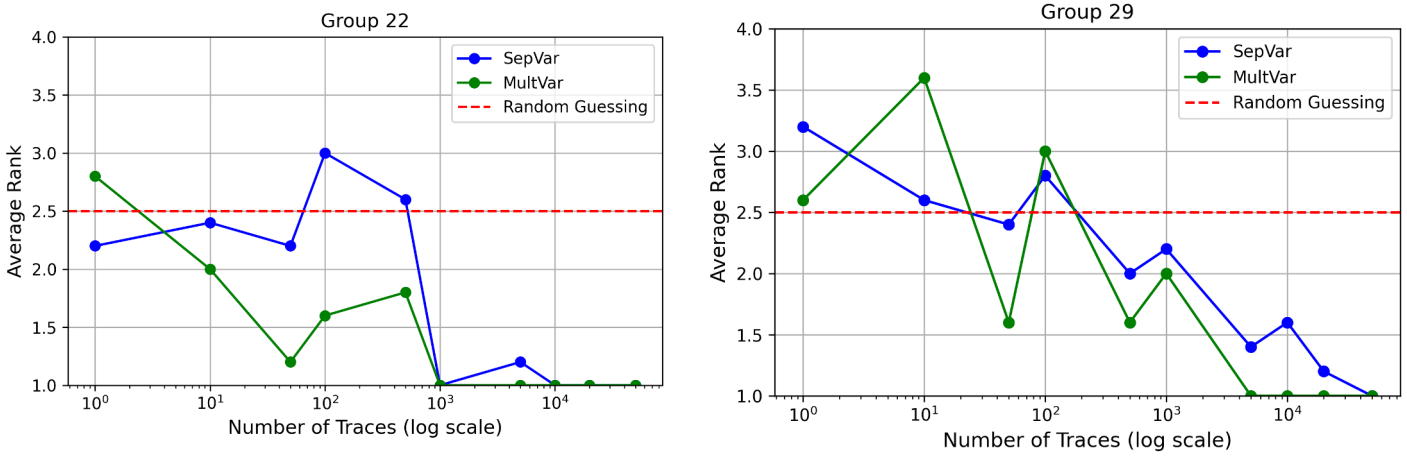


SepVar and MultVar - Accuracy over Epochs



SepVar and MultVar - Convergence Analysis

- Example 2-bit key groups convergence Analysis



Model	Converging Grps.	# of Power Traces	# of Parameters	Training Time (s)	Accuracy
SepVar	6/32	992	414,336	**7833**	0.0327
MultVar	**10/32**	**775**	**209,960**	13730	**0.0360**

Combined Results Table

- The table below summarizes the results from all six models.

Overview of the specifications and performance of the trained models

Model	Converging Grps.	# of Power Traces	# of Parameters	Training Time (s)	Accuracy
SepS	2/32	3050	414,336	**1147**	0.0311
MultS	6/32	592	**209,960**	3617	0.0349
SepL	0/32	N/A	414,336	2080	0.0318
MultL	**11/32**	**514**	**209,960**	7492	0.0357
SepVar	6/32	992	414,336	7833	0.0329
MultVar	10/32	775	**209,960**	13730	**0.0362**

SepVar and MultVar - Analysis Dataset

- It has been noted that the number of converging 2-bit key groups depends on the test dataset used.
- The analysis dataset was used to calculate the average performance of the variable-key models (SepVar and MultVar)

Dataset ID	Converging Grps.	
	SepVar	MultVar
0	1/32	3/32
1	1/32	2/32
2	4/32	10/32
3	5/32	11/32
4	2/32	8/32
5	2/32	4/32
6	1/32	5/32
7	5/32	9/32
8	1/32	7/32
9	6/32	10/32
μ	**2.8/32**	**6.9/32**
σ	**1.9889**	**3.2128**

Conclusion

- The **multi-head architecture** demonstrates much **better performance** when compared with their separate network model counterparts.
- **Variant success rates** when analyzing different power trace datasets.
- Future work could include comparing the results with **traditional attack methods** (e.g., CPA) and attacking **protected implementations** of the cipher.

Thanks for listening!

Any questions?